Bad return good return

Péter Erdős[§] PhD Student,
Department of Finance, Budapest University of Technology and Economics

Mihály Ormos[**] Associate Professor,
Department of Finance, Budapest University of Technology and Economics Budapest

Abstract
We study the methodology of return calculation in international mutual funds with investment in emerging markets particularly those with short market history. Besides daily and monthly returns we introduce a new procedure: the daily recalculated monthly returns (DRMR) and investigate which is the adequate method is for testing market efficiency and measuring fund performance. DRMR represents the log returns of the previous month (assuming 21 trading days in a month) and is recalculated day by day for the studied period. This method is effective for all assets and markets for which there is no suitably long time series available.

JEL classification numbers: F30; G12; G15
Keywords: Return calculation; International CAPM; Mutual fund performance

1. Introduction

We focus on return calculation when the main scope is performance measurement or testing market efficiency of debuting mutual funds in recently liberalized financial markets. Concentrating on these assets and markets, only short time series are available where annual or monthly returns cannot be used. We propose using a method for calculating the returns in empirical finance: we recalculate the monthly returns day by day assuming 21 trading days in each month which results in a daily recalculated monthly return (*DRMR*) time series. This method is widely used in the mutual fund industry for informing investors about the return of the previous month but not used in the academic literature. Using our procedure the number of observations is large enough to estimate a system of equations suitable for GRS test (Gibbons et. al, 1989) even when there is, for example, only a one-year-long time series available. The reliability of the parameters is better than that of resulting from monthly returns; while the goodness of fitting is better using DRMR comparing with simple monthly and daily returns.

Several authors apply monthly returns (e.g., Fama and MacBeth, 1973; Fama and French, 1992, 1995, 1996 a, b, 2006; Black et al., 1972; Elton et al., 1995, 1996) for testing the Capital Asset Pricing Model (CAPM) using at least five-year-long time series. Papers on mutual funds performance investigate whether portfolio managers can beat the market. The abnormal return can be determined by the alpha of the market model (Jensen, 1968). If the alpha is statistically not different from zero, the asset studied gains normal return. Several authors (e.g., Sharpe, 1966; Jensen, 1968; Kothari et al., 1995) use annual returns to measure decade-long mutual fund performance. These methods are adequate for a suitably long time series. However, if a shorter period is to be studied, the span horizon of returns is a key issue. If we use daily returns to obtain a long enough time series, the fitting of the market model deteriorates which means the standard errors of the parameters are greater. The univariate asset pricing tests are based on t-statistics which can sign no abnormal performance erroneously if the standard error of the regression is great. The GRS test is based on a

[§] email: erdos@finance.bme.hu
[**] email: ormos@finance.bme.hu

multivariate t-statistics therefore if the standard errors of the regressions are great the type II error can be unacceptable.

We apply International CAPM to compare the statistical properties of the estimations using daily, monthly and daily recalculated monthly returns when time series are short. We use Morgan Stanley Capital International (MSCI) All Countries World Index as a market proxy to measure the performance of the international mutual funds listed in Bloomberg.

Generally the GRS test and the Fama-MacBeth procedure are used for testing market efficiency. But these tests have strong statistical assumptions, as normally distributed, homoskedastic, and non-autocorrelated residuals are essential for the GRS test. If we use a short time series, the normality of monthly returns cannot be rejected. However, if we use longer time series, let's say at least 80 months simple monthly returns are not normal; on the other hand, the residuals of the market model still remain homoskedastic and non-autocorrelated. Assuming normality for a short time series (e.g., less than one-year) can be misleading because the Type II error is too great. We show that monthly and DRMRs have the same distribution since we can consider monthly returns as a systematic sample of DRMRs. We show that DRMRs are highly non-normal, which means monthly returns cannot be normal, either. Autocorrelation can be eliminated with systematic sampling, but the distribution is inherited in DRMRs.

If we know that the market model residuals do not behave like white noise, we still have the possibility to test efficiency or measure performance if the time series are long enough. We can apply asymptotic chi-square asset pricing tests under general conditions (see, e.g., Cochrane, 2001). If there is only a several-month-long monthly return time series available, we cannot derive appropriate asymptotic asset pricing tests. We can use daily returns; but as Morse (1984), Brown and Warner (1985) point out the fitting of the market model is better in the case of monthly returns. Fama (1976) also states that monthly returns are better working approximations than daily returns. In our data daily returns perform worse: the average R^2 is 55%, while in the case of monthly returns it is 62%. There are three different causes why we can see differences in fitting between monthly and daily returns as Morse (1984) shows: 1, there is a measurement error in the market proxy; 2, there is a measurement error in mutual fund's returns; 3, daily returns are not stationary.

We show that all the three types of returns are stationary according to ADF (Augmented Dickey-Fuller) tests (Dickey and Fuller, 1981) so the third story is not likely which means that the different goodness of fittings are attributable to measurement error. We show that DRMRs are less sensitive to measurement error either in the dependent variable or in the independent variable which is a preferable property because there is measurement error in the variables, otherwise the R^2s would be the same for daily and monthly returns. Higher R^2s means the squared error of the estimation is smaller thus inducing more reliable parameter estimations.

If we want to measure performance of debuting assets or test market efficiency of newly traded markets we propose the use of DRMRs especially if there is measurement error either in the assets' return or in the market proxy. We prefer the DRMRs to the two rival methods. Daily returns are more sensitive to measurement error resulting in low R^2s and monthly returns are applicable when time series are too short. The latter arises if you want to measure asset performance or test market efficiency by a system of equations (see e.g., Gibbons et al., 1989) and the number of assets multiplied by the number of parameters is larger than the number of observations.

In addition, there is also practical reason to support DRMRs. Monthly returns show the gain or loss over a particular month assuming an investment exactly at the beginning of the month. Investors are rather interested in actual monthly gain/loss which can be represented more precisely by DRMRs because they compare day-on-day (assuming 21 trading days in each month) not like monthly returns. If you use monthly returns, you may compare, for example, Friday with Monday. Measuring returns between different days of the week is not adequate because there is reason to assume that returns of different days show a different pattern (see e.g., Chen, Singal, 2003). Further, in the paper we are seeking for answers to the right choice of data resolution considering the pros and cons.

In Section II. we show that DRMRs induce the best fitting of the market model and they are the least sensitive to measurement errors either in the mutual fund's return or in the market proxy. In Section III. we show the invalidity of the assumptions of standard asset pricing tests. Daily and daily recalculated monthly returns are not normally distributed, heteroskedastic and serially dependent. Monthly returns seem to be normal only in small samples but they are a systematic sample of DRMR population. In the final section of the paper we summarize our findings.

2. Goodness of Fitting

Differences between daily and monthly returns are represented by the Single World Index Model of Solnik (1974). We use *GMM* (Generalized Method of Moments, Hansen, 1982) estimation testing the returns of 26 international open-ended mutual funds incorporated in the USA. Our data contain mutual funds listed in Bloomberg Mutual Fund Center and have at least a two-year market history; they exclude ETFs, REITs and institutional mutual funds. Analysis is carried out for the period from June 2003 to April 2008. Our mutual fund data are not free of survivorship bias (see e.g., Elton et. al,1996) but it does not affect our purpose to prove, DRMRs provide better working approximation than daily or monthly returns if the time series are short.

Solnik's model can be estimated as the CAPM with a global market portfolio,

$$R_{i,t} = \alpha_i + \beta_i (R_{m,t}) + \varepsilon_{i,t} \quad i = 1,2,...,N; \quad t = 1,2,...,T \quad , \qquad (1)$$

where $R_{i,t}$ and $R_{m,t}$ are the mutual fund risk premium and the market risk premium respectively. Continuously compounded returns are calculated in US dollars.

We use Morgan Stanley Capital International (MSCI) All Countries World Index as the market proxy and the one-month US Treasury Bill returns by Ibbotson Associates as the risk-free rate. Fama and French (1998) use their own world index containing the components of MSCI EAFE (Europe, Asia and Far East) Index and the MSCI USA index additionally for international asset pricing. De Santis and Gerard (1997) use MSCI World Index for international asset pricing of developed countries. The investment focus of the mutual funds in our data is not restricted to the EAFE+US countries so we choose a broader index which includes emerging markets as well. The regional exposure and the style of each mutual fund can be seen in Table 1.

Regional exposures and styles

Table 1

Panel A: Mutual fund number 1-13

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
Regional	Greater	China	Greater	Greater	Greater	Russia	Russia	Korea	UK	EM	EM	Global	Global
Style	LB	LV	LB	LG	LG	N/A	N/A	LB	SB	LB	LB	MG	LB

Panel B: Mutual fund number 13-26

	14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
Regional	Global	Global	Global	Global	Global	Global	Europe	Greater	Latin	Latin	Nordic	Southeast	Asia ex-
Style	LV	LG	LB	LG	MB	LG	LB	LG	LB	LG	LG	LV	LG

This table shows regional exposures and styles of mutual funds included in the data set. EM stands for Emerging Markets. L, M, S, stand for large, mid and small cap, and V, B, G indicates the investment style, value, blend and growth respectively using Morningstar's database.

As table 2 shows the goodness of fitting in the case of monthly returns measured by R^2 is seven percentage points higher than that of the daily returns (We have to note that R^2 is not suitable for direct comparisons if the dependent variables are different.We use it for illustrative purpose as Morse (1984), and Warner and Brown (1985).). The average R^2 calculated on the basis of daily returns is 55%, which is lower than the 62% statistics of monthly returns. Monthly returns have better R^2s in 21 cases out of 26. Significant improvement can be observed in some cases if monthly return is used instead of daily, particularly in mutual fund number 2, 21, 26. The deterioration is less significant in five cases. Average R^2s are relatively high compared with the results of Fama and French (1998) for country portfolios of 12 EAFE countries plus the US. Using monthly returns their average R^2 is only 39%.

If conditions of asset pricing tests (normality, homoskedasticity, serial independence) are not met, we propose the use of DRMRs which are more appropriate for large sample tests if the time series is not long enough, and the fitting of the market model is even better than in the case of monthly returns.

International CAPM estimated by GMM regressions using daily, monthly and daily recalculated monthly returns

Table 2

$$R_{i,t} = \alpha_i + \beta_i (R_{m,t}) + \varepsilon_{i,t} \quad i = 1,2,...,26; \quad t = 1,2,...,T$$

No.	A					B					C				
	α	$t(\alpha)$	β	$t(\beta)$	R^2	α	$t(\alpha)$	β	$t(\beta)$	R^2	α	$t(\alpha)$	β	$t(\beta)$	R^2
1	0.06	1.43	1.27	18.54	0.51	1.29	1.09	1.55	5.20	0.55	0.69	0.81	1.52	9.55	0.69
2	0.04	1.33	0.84	14.16	0.25	0.28	0.48	1.53	7.08	0.53	0.07	0.15	1.38	12.32	0.63
3	0.04	1.13	1.30	20.37	0.39	0.60	0.77	1.59	6.55	0.47	0.22	0.30	1.59	9.69	0.59
4	0.03	1.19	1.15	21.12	0.39	0.36	0.56	1.67	7.69	0.56	0.10	0.16	1.52	10.39	0.61
5	0.08	1.69	1.35	16.59	0.48	1.54	1.80	1.77	5.66	0.61	0.96	0.99	1.77	11.15	0.71
6	0.07	1.71	1.13	16.80	0.24	1.12	1.37	1.44	3.64	0.29	0.70	0.61	1.46	5.82	0.37
7	0.06	1.76	0.83	15.90	0.21	0.91	1.09	1.27	3.73	0.26	0.60	0.49	1.28	5.13	0.35
8	0.02	0.74	1.16	22.67	0.36	0.27	0.53	1.37	9.09	0.48	0.03	0.05	1.29	9.53	0.52
9	0.02	1.06	0.96	25.25	0.51	0.33	0.82	1.04	5.42	0.45	-0.02	-0.03	1.05	8.56	0.57
10	0.04	2.41	1.37	36.48	0.70	0.71	1.56	1.58	7.89	0.73	0.35	0.67	1.56	9.53	0.77
11	0.04	2.25	1.37	36.81	0.70	0.65	1.41	1.58	7.87	0.73	0.30	0.57	1.56	9.49	0.77
12	-0.02	-0.71	1.61	34.94	0.54	0.10	0.20	1.37	7.48	0.44	-0.34	-0.56	1.49	10.20	0.59
13	0.03	4.60	0.59	47.42	0.77	0.48	4.05	0.67	16.52	0.82	0.32	2.22	0.63	15.86	0.83
14	-0.01	-0.96	1.09	45.34	0.77	0.07	0.53	0.87	13.31	0.75	-0.24	-1.41	0.99	13.97	0.85
15	0.01	2.00	1.03	76.13	0.88	0.30	1.97	1.02	19.63	0.89	0.05	0.33	1.00	22.96	0.91
16	0.01	1.14	0.93	60.21	0.82	0.11	0.66	1.05	20.12	0.83	-0.11	-0.52	1.02	19.10	0.86
17	0.00	0.09	1.24	65.67	0.83	0.21	1.04	1.04	16.82	0.84	-0.18	-1.39	1.15	26.24	0.90
18	0.05	3.53	1.09	31.39	0.66	0.87	2.52	1.39	12.26	0.70	0.53	1.31	1.36	12.96	0.77
19	0.02	1.55	1.33	48.42	0.75	0.48	2.20	1.20	16.27	0.75	0.07	0.33	1.32	21.13	0.85
20	0.01	0.88	1.22	42.91	0.68	0.23	0.84	1.27	13.64	0.70	-0.06	-0.19	1.24	13.48	0.76
21	0.08	1.85	0.99	13.45	0.20	0.65	0.61	2.13	7.64	0.50	0.63	0.59	1.72	6.11	0.47
22	0.07	2.47	1.79	32.28	0.58	1.49	2.54	1.65	6.57	0.55	0.91	1.58	1.80	10.52	0.70
23	0.05	1.80	1.80	32.29	0.56	1.19	2.03	1.58	6.27	0.53	0.64	1.14	1.74	10.73	0.68
24	0.02	1.25	1.31	35.59	0.63	0.37	1.25	1.46	15.55	0.75	0.12	0.45	1.37	24.80	0.80
25	0.04	1.62	1.19	23.99	0.46	0.56	0.97	1.62	9.33	0.64	0.24	0.41	1.57	9.99	0.68
26	0.04	1.89	1.04	26.66	0.48	0.53	1.56	1.38	13.06	0.70	0.22	0.65	1.34	15.96	0.74
Ave	**0.03**		**1.19**		**0.55**	**0.60**		**1.39**		**0.62**	**0.26**		**1.37**		**0.69**

This table shows the alphas, betas and their t-statistics of the market model using daily, monthly and daily recalculated monthly returns in panel A, B, C respectively. We use the Morgan Stanley Capital International (MSCI) All Countries World Index as the market proxy. $t(\alpha)$ and $t(\beta)$ are t-statistics with Heteroskedasticity and Autocorrelation Consistent (HAC) standard errors (Andrews 1991). R^2s are not adjusted with degrees of freedom.

Figure 1 shows the characteristic lines of monthly and daily returns of the same mutual fund in Panel A and Panel B respectively. Fitting is better in the case of monthly returns because they spread around the regression line randomly, on the other hand, daily returns concentrate.

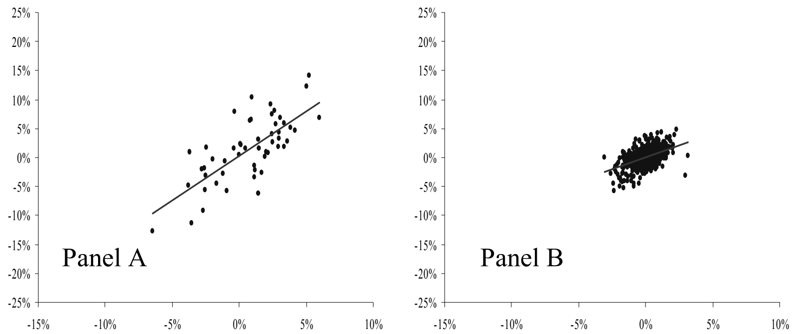

Figure 1 Characteristic line ($R_{2,t}=\alpha_2+\beta_2(R_{m,t})+\varepsilon_{2,t}$) of mutual fund number 2.

Legend: x axis is $R_{m,t}$ (market risk premium), y axis is $R_{2,t}$ (risk premium of mutual fund number 2). Notes: Characteristic line ($R_{2,t}=\alpha_2+\beta_2(R_{m,t})+\varepsilon_{2,t}$) of mutual fund number 2 based on monthly and daily return in Panel A and B respectively. Mutual fund number 2 is chosen for illustration because it has much better fitting when monthly returns are used instead of daily.

Comparing Table 2 panel C with regressions on monthly returns (Table 2 panel B), we can see that the estimated betas are approximately the same. The average R^2 is significantly better than in the case of daily and monthly returns. In 20 cases out of 21 (monthly returns in 21 cases were better than daily ones) further improvement can be recognized if we use DRMRs instead of monthly returns. In four out of five cases when monthly returns fail to be better than daily ones, DRMRs do a better job than monthly returns. There is only one mutual fund in our data set which produces higher R^2 with monthly returns, although the difference is only 3 percentage points. DRMRs' fitting is significantly better than daily returns' and even better than that of the monthly returns'.

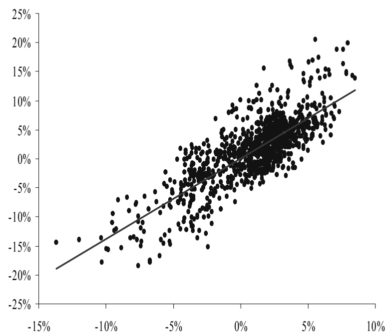

Figure 2 Characteristic line ($R_{19,t}=\alpha_{19}+\beta_{19}(R_{m,t})+\varepsilon_{19,t}$) of mutual fund number 19 using daily recalculated monthly returns (*DRMR*s)

Legend: x axis is $R_{m,t}$ (market risk premium), y axis is $R_{19,t}$ (risk premium of mutual fund number 19). Notes: Mutual fund number 19 is chosen for illustration because its R2 is equal to the average of the 26 market model regressions included in our paper when using DRMRs.

Let us assume that only a one-year-long time series is available and the performance of a specific asset is to be measured. If we use simple monthly returns only twelve data points would be available for the estimation. It does not remain without consequences: goodness of fitting decreases dramatically. As an illustration, we choose one mutual fund from our data set and compare fitting of DRMRs with monthly returns.

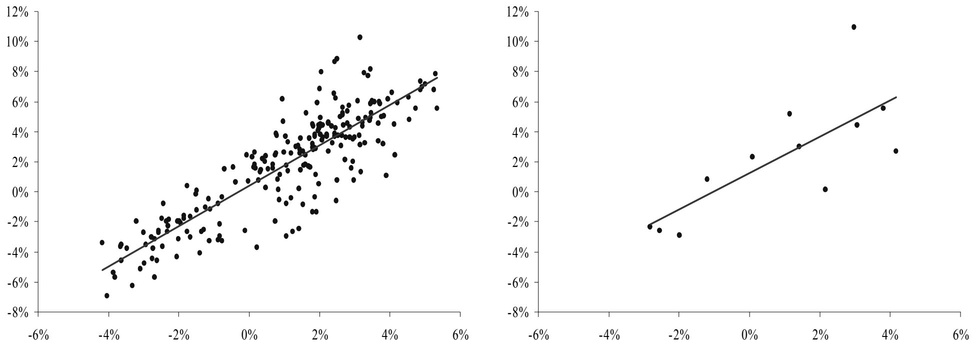

Figure 3 Characteristic line *($R_{19,t}=\alpha_{19}+\beta_{19}(R_{m,t})+\varepsilon_{19,t}$)* of mutual fund number 19

Legend: *x* axis is $R_{m,t}$ (market risk premium), *y* axis is $R_{19,t}$ (risk premium of mutual fund number 19)
Notes: Figure shows characteristic lines ($R_{19,t}=\alpha_{19}+\beta_{19}(R_{m,t})+\varepsilon_{19,t}$) of mutual fund number 19 in the case of daily recalculated monthly returns (*DRMR*s) on Panel A, and in the case of monthly returns on Panel B for the year 2005. Mutual fund number 19 is chosen for illustration because its $\boldsymbol{R^2}$ is equal to the average of the 26 market model regressions included in our paper when using *DRMR*s.

On Figure 3, Panel A the data points are spread in a well defined band around the regression line, therefore fitting is adequate. In the case of monthly returns (Panel B) the fitting is worse.

Running the CAPM regression produces similar fitting in both monthly methods which means that recalculating the monthly returns daily gives at least as good fitting as we can see in the case of simple monthly returns, however much shorter time interval is enough for an adequate analysis.

There can be different causes of the inequality of R^2s of monthly and daily returns. If there is no measurement error and the daily returns are stationary the R^2s would be the same independently of the method of return calculation (see Morse, 1984 for the proof). We assume that there is measurement error at least in the market proxy because of imperfect market portfolio (see e.g., Roll, 1977). DRMRs, daily and monthly returns are stationary at any usual significance level based on ADF tests.

We test the effect of deliberately added random errors to the mutual funds' returns and to the market proxy using the Monte Carlo experiment. We find that DRMRs are the least and daily returns are the most sensitive to measurement errors. We generate numbers from Student t-distribution with four degrees of freedom. Then we estimate the following models:

$$\begin{aligned} &a, \quad R_{i,t} = \alpha_i' + \beta_i'\left(R_{m,t} + c\tau_t\right) + \varepsilon_{i,t}' \quad i = 1,2,...,N; \quad t = 1,2,...,T \\ &b, \quad \left(R_{i,t} + c'\tau_t'\right) = \alpha_i'' + \beta_i'' R_{m,t} + \varepsilon_{i,t}'' \quad i = 1,2,...,N; \quad t = 1,2,...,T \end{aligned}, \quad (2)$$

where τ is the generated random error from t-distribution, and c is a constant which ensures the variance of the new variable to be (market proxy/mutual funds' returns plus the generated error) 1.2, 1.4,…,2 times more than the original variable. We use t-distribution with four degrees of freedom because it makes the distribution of the (in)dependent variable fat-tailed, which is a typical feature of asset returns. We choose four degrees of freedom to ensure to be the first four moments of the distribution finite.

Table 3 shows that DRMRs are the least sensitive to measurement error. If we use DRMRs and add random errors to the market proxy (mutual fund's returns) which makes the variance 1.2 times the original, the market model's R^2 is 45.22% (57.58%) on average.

Adding random errors to daily returns the average R^2 is 37.19% (46.19%) and to monthly returns it is 39.57% (51.31%).

The effect of random errors on the market proxy and the mutual funds' returns

Table 3

	Panel A						Panel B					
	1.0	1.2	1.4	1.6	1.8	2.0	1.0	1.2	1.4	1.6	1.8	2.0
	Daily returns											
1	51.38	31.69	22.98	18.04	14.86	12.65	51.38	43.06	37.11	32.62	29.12	26.30
2	24.89	15.70	11.51	9.09	7.51	6.40	24.89	20.83	17.92	15.74	14.03	12.67
3	38.67	21.05	14.51	11.08	8.98	7.55	38.67	32.51	28.07	24.70	22.06	19.94
4	39.33	23.69	16.99	13.26	10.87	9.22	39.33	33.02	28.47	25.04	22.35	20.19
5	47.57	27.30	19.21	14.85	12.12	10.25	47.57	39.83	34.31	30.16	26.92	24.32
6	24.19	12.22	8.21	6.19	4.98	4.17	24.19	20.44	17.69	15.61	13.96	12.64
7	21.27	13.14	9.54	7.50	6.18	5.27	21.27	17.88	15.43	13.58	12.13	10.96
8	35.61	19.66	13.62	10.43	8.46	7.12	35.61	29.41	25.05	21.83	19.35	17.38
9	51.06	38.11	30.44	25.35	21.73	19.02	51.06	42.96	37.11	32.67	29.19	26.39
10	69.60	45.83	34.22	27.32	22.76	19.51	69.60	58.42	50.37	44.29	39.52	35.69
11	69.74	45.99	34.36	27.45	22.87	19.61	69.74	58.55	50.49	44.40	39.62	35.78
12	53.90	27.99	18.96	14.36	11.56	9.68	53.90	45.39	39.15	34.43	30.73	27.75
13	77.37	71.14	65.82	61.25	57.28	53.79	77.37	64.63	55.52	48.68	43.35	39.08
14	77.45	58.70	47.30	39.63	34.12	29.96	77.45	64.23	54.87	47.90	42.52	38.22
15	88.02	70.88	59.36	51.08	44.84	39.97	88.02	73.44	63.02	55.21	49.13	44.26
16	81.65	67.09	57.00	49.57	43.86	39.33	81.65	67.91	58.16	50.87	45.22	40.70
17	83.31	60.61	47.67	39.32	33.47	29.14	83.31	69.30	59.32	51.86	46.09	41.47
18	65.53	49.06	39.22	32.69	28.04	24.56	65.53	55.17	47.68	42.00	37.53	33.94
19	75.10	51.24	38.93	31.42	26.36	22.71	75.10	62.78	53.94	47.30	42.12	37.97
20	68.03	47.20	36.17	29.35	24.71	21.35	68.03	56.44	48.28	42.20	37.49	33.73
21	20.46	10.93	7.49	5.71	4.63	3.90	20.46	17.30	14.98	13.22	11.83	10.70
22	57.93	28.25	18.74	14.04	11.23	9.37	57.93	48.79	42.14	37.09	33.13	29.94
23	56.40	26.98	17.79	13.29	10.62	8.85	56.40	47.53	41.06	36.14	32.28	29.18
24	62.84	40.86	30.32	24.13	20.04	17.15	62.84	52.49	45.10	39.55	35.23	31.76
25	46.10	28.39	20.55	16.12	13.27	11.28	46.10	38.42	32.97	28.88	25.70	23.16
26	48.21	33.23	25.39	20.56	17.28	14.91	48.21	40.21	34.49	30.21	26.88	24.22
Average	55.22	37.19	28.70	23.58	20.10	17.57	55.22	46.19	39.72	34.85	31.06	28.01

	Panel A						Panel B					
	1.0	1.2	1.4	1.6	1.8	2.0	1.0	1.2	1.4	1.6	1.8	2.0
	Monthly returns											
1	55.26	31.61	23.02	18.51	15.74	13.85	55.26	47.47	41.69	37.30	33.85	31.06
2	53.48	30.09	21.40	16.83	14.01	12.09	53.48	45.20	39.23	34.73	31.21	28.38
3	46.96	23.46	16.14	12.51	10.34	8.90	46.96	39.52	34.25	30.28	27.18	24.69
4	55.78	28.47	19.66	15.25	12.59	10.81	55.78	46.48	40.07	35.29	31.58	28.61
5	61.06	32.14	22.74	18.02	15.18	13.27	61.06	51.88	45.43	40.59	36.79	33.73
6	28.92	13.31	9.11	7.16	6.03	5.29	28.92	24.28	21.04	18.62	16.73	15.22
7	26.45	13.20	9.20	7.28	6.15	5.40	26.45	22.25	19.28	17.06	15.33	13.94
8	47.73	24.22	16.54	12.69	10.38	8.83	47.73	38.33	32.19	27.83	24.56	22.02
9	45.06	29.75	22.59	18.39	15.62	13.65	45.06	37.24	31.82	27.83	24.78	22.36
10	73.01	44.73	32.66	25.94	21.64	18.66	73.01	61.23	52.80	46.48	41.56	37.62
11	73.11	44.68	32.59	25.86	21.58	18.60	73.11	61.32	52.88	46.55	41.63	37.68
12	44.04	24.97	17.85	14.11	11.81	10.25	44.04	37.97	33.13	29.40	26.45	24.06
13	81.52	72.17	64.79	58.83	53.92	49.80	81.52	66.24	55.79	48.27	42.59	38.15
14	74.82	62.00	52.96	46.34	41.29	37.29	74.82	61.53	52.33	45.60	40.46	36.40
15	88.58	71.92	60.64	52.54	46.42	41.64	88.58	73.47	63.00	55.23	49.23	44.45
16	83.00	63.99	52.29	44.31	38.52	34.11	83.00	67.77	57.46	49.98	44.29	39.81
17	84.40	66.84	55.47	47.54	41.69	37.18	84.40	69.70	59.35	51.74	45.92	41.32
18	70.32	46.96	35.59	28.85	24.38	21.19	70.32	59.07	51.08	45.07	40.38	36.61
19	74.97	54.06	42.48	35.17	30.11	26.40	74.97	61.51	52.46	45.84	40.78	36.76
20	70.22	46.50	35.17	28.46	24.01	20.84	70.22	57.03	48.34	42.06	37.28	33.53
21	49.82	18.57	11.73	8.71	7.02	5.94	49.82	42.13	36.61	32.44	29.16	26.52
22	55.39	30.31	21.38	16.83	14.05	12.18	55.39	46.56	40.14	35.33	31.58	28.59
23	52.80	29.59	21.04	16.62	13.91	12.08	52.80	44.36	38.25	33.67	30.11	27.26
24	74.72	45.84	33.46	26.55	22.13	19.06	74.72	60.89	51.57	44.82	39.69	35.65
25	63.95	35.49	24.99	19.46	16.05	13.74	63.95	53.22	45.77	40.23	35.94	32.52
26	69.66	43.97	32.49	25.91	21.63	18.63	69.66	57.41	49.03	42.87	38.14	34.40
Average	61.73	39.57	30.31	24.95	21.39	18.83	61.73	51.31	44.04	38.66	34.51	31.21

	Panel A						Panel B					
	1.0	1.2	1.4	1.6	1.8	2.0	1.0	1.2	1.4	1.6	1.8	2.0
	Daily recalculated monthly returns											
1	68.81	41.47	29.82	23.35	19.21	16.35	68.81	57.61	49.59	43.55	38.84	35.05
2	62.90	39.27	28.59	22.50	18.56	15.81	62.90	52.69	45.33	39.78	35.45	31.97
3	58.76	32.04	22.07	16.84	13.63	11.45	58.76	49.42	42.63	37.48	33.45	30.21
4	61.15	34.86	24.42	18.79	15.28	12.88	61.15	51.28	44.15	38.77	34.57	31.19
5	71.27	38.22	26.27	20.07	16.28	13.72	71.27	59.77	51.47	45.21	40.32	36.39
6	36.87	17.40	11.43	8.53	6.81	5.68	36.87	31.04	26.79	23.56	21.04	19.01
7	34.64	18.05	12.24	9.27	7.48	6.27	34.64	29.17	25.17	22.14	19.76	17.85
8	52.23	30.21	21.29	16.46	13.42	11.34	52.23	42.69	36.11	31.29	27.61	24.71
9	57.33	40.02	30.80	25.03	21.08	18.21	57.33	47.13	40.03	34.80	30.78	27.60
10	77.20	47.68	34.53	27.09	22.30	18.96	77.20	64.76	55.78	49.00	43.70	39.43
11	77.27	47.73	34.57	27.12	22.33	18.98	77.27	64.82	55.83	49.04	43.73	39.47
12	58.95	33.85	23.78	18.35	14.94	12.61	58.95	49.48	42.63	37.46	33.41	30.16
13	83.46	74.91	67.97	62.22	57.37	53.22	83.46	67.88	57.23	49.48	43.59	38.96
14	85.27	68.26	56.94	48.85	42.78	38.06	85.27	70.37	59.89	52.13	46.16	41.42
15	90.50	73.46	61.86	53.43	47.04	42.01	90.50	75.12	64.24	56.13	49.85	44.84
16	86.30	68.21	56.42	48.12	41.95	37.20	86.30	71.05	60.39	52.53	46.48	41.69
17	90.20	68.83	55.69	46.78	40.34	35.46	90.20	74.77	63.85	55.73	49.45	44.44
18	77.40	52.79	40.09	32.33	27.10	23.33	77.40	65.00	56.04	49.27	43.97	39.70
19	85.27	59.94	46.25	37.67	31.79	27.50	85.27	70.85	60.64	53.02	47.11	42.39
20	75.81	52.67	40.39	32.77	27.58	23.81	75.81	62.63	53.37	46.51	41.22	37.01
21	47.30	21.75	14.16	10.52	8.38	6.97	47.30	39.97	34.59	30.49	27.27	24.66
22	70.13	36.71	24.92	18.88	15.22	12.75	70.13	58.67	50.45	44.26	39.44	35.56
23	67.92	36.12	24.65	18.73	15.12	12.68	67.92	56.91	48.96	42.98	38.30	34.54
24	79.57	52.41	39.12	31.22	25.99	22.26	79.57	65.49	55.68	48.45	42.89	38.48
25	67.86	39.46	27.86	21.55	17.58	14.85	67.86	56.82	48.87	42.89	38.21	34.46
26	74.00	49.31	37.02	29.65	24.73	21.22	74.00	61.63	52.81	46.20	41.07	36.97
Average	69.17	45.22	34.35	27.93	23.63	20.52	69.17	57.58	49.33	43.16	38.37	34.55

This table shows the R^2s of the market model fitting on each mutual fund's daily, monthly and daily recalculated monthly returns. We add random errors generated from t distribution with four degrees of freedom to the market proxy in Panel A and to mutual funds's return in Panel B. We multiply the errors with a suitable scalar in each column to obtain a market/mutual fund premium which has exactly 1, 1.2,…,2 times the variance the original variable has. The R2s are the average of 100 repeats Monte Carlo experiment.

3. Invalidity of the assumptions of standard asset pricing tests

Small and large sample tests must be differentiated in empirical studies. Distribution of small sample test statistics can usually be derived only under strong assumptions, such as normality and other specific ones like homoscedastic, or non-autocorrelated error terms.

Large sample tests are valid under less restrictive conditions but sufficiently large samples are not always available.

If we want to apply the GRS test for testing market efficiency which is usually used in the literature the following assumptions should be satisfied:

1. $\varepsilon_{i,t} \left| R_{m,s} \sim N\left(0, \Sigma = I\sigma^2(\varepsilon)\right);\right. \; t,s = 1,2,...,T$, i.e. the conditional distribution of the residuals is normal with constant standard deviation (homoscedastic) and zero conditional expected value. $I\left(N \times N\right)$ is the identity matrix, $\sigma^2(\varepsilon)$ is the residual variance, $\Sigma = E\left(\varepsilon_t \varepsilon_t'\right)$, where $\varepsilon_t' = \left(\varepsilon_{1,t}, \varepsilon_{2,t}, ..., \varepsilon_{N,t}\right)$.
2. $E\left(\varepsilon_t \varepsilon_{t-l}'\right) = 0; \; l = 1,2,...,t+1$ (non-autocorrelated).

If we use a small sample test like GRS, we have to test the validity of the above conditions. The Fama-MacBeth (1973) procedure also assumes normality in the residuals. One of the main difficulties of testing normality of small samples by exact normality tests (e.g., D'Agostino-Pearson K^2, Jarque-Bera, 1980, etc.) is that normality of a suitably small sample can almost be rejected. If we study a one-year-long time series and calculate monthly returns, we have only 12 data points available for testing normality which can be misleading because of Type II error. We use D'Agostino-Pearson K^2 statistics for testing normality (see e.g., D'Agostino et al., 1990), which is a statistically strong test even on small samples (more than 9 elements). We apply Jarque-Bera normality test as well, the results are the same as in the case of D'Agostino-Pearson K^2 tests.

Table 4 shows that daily returns are highly nonnormal (26 time series out of 26) while monthly returns appear to be normal (25 time series out of 26) for a five-year-long period. If we use a longer period of at least 80 months (there are 23 mutual funds with such a long market history in our data) normality tests reject the normal null hypothesis at 95% significance level (see Table 5). None of the DRMR time series are normal at any usual significance level.

K^2 chi-square test is a sum of squares of two separate normally distributed test statistics for skewness and kurtosis known as directional tests. We find that there are only two DRMR time series which exhibit normal distribution based on kurtosis (see Table 4).

Normality tests based on skewness and kurtosis and D'Agostino-Pearson K^2 test of return data

Table 4

Panel A

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
Z(1)	-3,73	-6,18	-4,14	-7,07	-5,90	-10,60	-10,26	-4,20	-8,12	-7,54	-7,65	-5,35	-5,74
p(Z(1))	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Z(2)	3,61	9,44	9,60	8,98	4,90	12,61	11,10	6,81	9,42	8,70	8,72	3,64	6,26
p(Z(2))	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
K^2	26,93	127,44	109,35	130,58	58,74	271,46	228,38	64,02	154,73	132,48	134,53	41,89	72,19
$p(K^2)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

	14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
Z(1)	-5,09	-6,13	-6,08	-1,91	-9,67	-6,76	-8,09	-9,61	-6,31	-5,80	-7,92	-8,00	-6,46
p(Z(1))	0,00	0,00	0,00	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Z(2)	6,16	6,72	5,56	8,30	10,52	7,84	9,27	12,73	8,20	8,13	8,98	9,39	7,99
p(Z(2))	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
K^2	63,93	82,69	67,91	72,62	204,13	107,10	151,49	254,31	107,02	99,85	143,43	152,24	105,64
$p(K^2)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Panel B

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
Z(1)	-0.50	-0.80	-0.27	-0.71	-1.18	-1.61	-1.44	0.40	-1.40	-1.88	-1.89	-0.73	0.38
p(Z(1))	0.35	0.29	0.38	0.31	0.20	0.11	0.14	0.37	0.15	0.07	0.07	0.31	0.37
Z(2)	-0.10	0.67	0.13	0.81	1.01	-0.01	0.08	0.66	1.44	0.16	0.16	0.13	0.80
p(Z(2))	0.40	0.32	0.40	0.29	0.24	0.40	0.40	0.32	0.14	0.39	0.39	0.40	0.29
K^2	0.26	1.09	0.09	1.17	2.41	2.59	2.09	0.60	4.02	3.57	3.59	0.55	0.79
$p(K^2)$	0.88	0.58	0.96	0.56	0.30	0.27	0.35	0.74	0.13	0.17	0.17	0.76	0.67

	14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
Z(1)	0.21	-1.13	-1.00	-0.15	-1.66	-0.44	-0.61	-3.02	-1.18	-1.19	-0.35	-1.08	-0.51
p(Z(1))	0.39	0.21	0.24	0.39	0.10	0.36	0.33	0.00	0.20	0.20	0.38	0.22	0.35
Z(2)	-1.63	-0.56	-0.44	-1.07	0.62	-0.09	1.36	1.75	0.88	0.83	-0.16	-0.46	0.12
p(Z(2))	0.11	0.34	0.36	0.22	0.33	0.40	0.16	0.09	0.27	0.28	0.39	0.36	0.40
K^2	2.69	1.60	1.18	1.17	3.12	0.20	2.21	12.19	2.16	2.11	0.15	1.36	0.27
$p(K^2)$	0.26	0.45	0.55	0.56	0.21	0.90	0.33	0.00	0.34	0.35	0.93	0.51	0.87

Panel C

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
Z(1)	-3.26	-5.08	-5.38	-7.39	-5.64	-10.69	-9.36	-8.93	-7.18	-13.80	-13.78	-6.03	-7.39
p(Z(1))	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Z(2)	1.44	5.61	6.36	5.66	2.02	2.82	2.50	3.43	2.38	8.47	8.44	1.28	4.35
p(Z(2))	0.14	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.02	0.00	0.00	0.18	0.00
K^2	12.70	57.31	69.35	86.58	35.91	122.34	93.91	91.49	57.19	262.21	261.10	38.00	73.57
$p(K^2)$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
Z(1)	-9.44	-10.85	-11.28	-10.13	-12.51	-10.78	-10.56	-15.52	-13.03	-12.23	-9.02	-11.05	-9.15
p(Z(1))	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Z(2)	5.22	4.87	4.55	5.40	6.60	5.44	6.00	9.31	8.22	7.62	2.08	6.63	4.73
p(Z(2))	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00
K^2	116.37	141.34	147.99	131.75	200.04	145.68	147.46	327.64	237.27	207.72	85.66	166.01	106.09
$p(K^2)$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

This table shows the test statistics of normality tests and their p-values of daily, monthly and daily recalculated monthly returns in panel A, B, C respectively. Z(1) tests normality with an alternative hypothesis of non-normal distribution due to skewness. Z(2) tests normality with an alternative hypothesis of non-normal distribution due to kurtosis. D'Agostino-Pearson K2 test statistic is the sum of squares of Z(1) and Z(2) statistics, which tests normality against non-normal distribution due to skewness and kurtosis.

Normality tests of monthly returns based on full history

Table 5

	1	2	3	4	5	6	7	8	9	10	11	12	13
Obs	28.00	54.00	129.00	102.00	28.00	137.00	91.00	140.00	177.00	143.00	143.00	269.00	190.00
Z(1)	1.05	2.82	6.51	5.03	0.80	5.18	2.19	4.21	5.90	6.12	6.10	4.69	5.03
p(Z(1))	0.23	0.01	0.00	0.00	0.29	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00
Z(2)	-1.92	0.13	3.86	2.29	-1.67	1.19	-3.84	-0.60	1.69	2.59	2.58	0.00	-0.64
p(Z(2))	0.06	0.40	0.00	0.03	0.10	0.20	0.00	0.33	0.10	0.01	0.01	0.40	0.32
K^2	4.79	7.98	57.31	30.51	3.42	28.29	19.52	18.13	37.64	44.12	43.86	21.98	25.67
$p(K^2)$	0.09	0.02	0.00	0.00	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	14	15	16	17	18	19	20	21	22	23	24	25	26
Obs	143.00	143.00	135.00	143.00	89.00	82.00	134.00	143.00	143.00	83.00	136.00	179.00	143.00
Z(1)	2.50	4.49	2.72	4.67	2.38	2.66	4.97	5.87	6.35	3.21	4.64	8.01	5.14
p(Z(1))	0.02	0.00	0.01	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Z(2)	-2.14	0.54	-3.22	1.00	-4.17	-2.27	0.78	2.44	2.92	-0.84	0.82	4.87	1.53
p(Z(2))	0.04	0.35	0.00	0.24	0.00	0.03	0.29	0.02	0.01	0.28	0.29	0.00	0.12
K^2	10.84	20.47	17.77	22.80	23.05	12.20	25.27	40.45	48.85	11.00	22.22	87.85	28.77
$p(K^2)$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

This table shows the test statistics of normality tests of monthly returns and their p-values. Full market history was used for testing normality in each case. Obs is the number of observation. Z(1) test of normality with alternative hypothesis of non-normal distribution due to skewness. Z(2) test of normality with alternative hypothesis of non-normal distribution due to kurtosis. D'Agostino-Pearson K2 test statistic is the sum of squares of Z(1) and Z(2) statistics, which tests normality against non-normal distribution due to skewness and kurtosis.

We also perform studentized range (SR) normality test based on the distribution of studentized range of standard normal random variable. It is not easy to define the distribution of range-to-deviation ratio but its quantiles can be reproduced by the Monte Carlo simulation (see Table 6, panel B). The method we use is the same as in Fama (1976). SR statistics lead to the same results as the K^2 and Jarque-Bera tests in the case of monthly returns. Table 6 panel A shows that the results are contradictory to our previous results applying SRs on daily returns and on DRMRs. Daily returns are non-normal variables, although the SRs show normality three times. DRMRs appear to be normal based on SR statistics (24 time series out of 26).

Studentized ratios of mutual funds

Table 6

Panel A

	1	2	3	4	5	6	7	8	9	10	11	12	13
Daily	7.29	**8.82**	**9.10**	**9.93**	7.03	**11.91**	**9.65**	**9.15**	**8.78**	**8.78**	**8.78**	6.80	**8.17**
Daily recalculated monthly	6.29	6.87	7.14	**7.69**	**5.57**	6.45	6.45	6.37	6.38	6.73	6.72	6.54	6.65
Monthly	4.06	4.83	4.76	4.85	4.61	4.68	4.61	5.20	5.55	4.41	4.41	5.06	4.95

	14	15	16	17	18	19	20	21	22	23	24	25	26
Daily	**7.75**	**8.51**	**8.27**	**8.55**	**9.73**	**8.74**	**9.10**	**11.88**	**8.08**	**9.00**	**9.09**	**9.81**	**9.23**
Daily recalculated monthly	7.36	6.24	5.89	6.81	6.54	6.66	6.79	7.08	6.92	6.91	5.93	7.24	6.61
Monthly	4.14	4.34	4.55	4.06	5.01	4.92	**5.86**	4.83	4.66	4.756	4.86	4.623	4.743

Panel B

	0.05%	0.5%	2.5%	5%	10%	90%	95%	97.5%	99.5%	99.95%
Normal sample (1218)	5.45	5.64	5.81	5.91	6.04	7.22	7.44	7.65	8.07	8.63
Normal sample (1198)	5.44	5.61	5.79	5.90	6.02	7.22	7.44	7.66	8.10	8.63
Normal sample (58)	3.46	3.64	3.82	3.92	4.04	5.23	5.44	5.63	6.04	6.53

This table shows the empirical studentized ranges of daily recalculated monthly and monthly returns in Panel A. Our data sets contain 1218 (in the case of daily returns), 1198 (in case of daily recalculated monthly returns) and 58 (in the case of monthly returns) observations. Panel B shows the percentiles of studentized range of standard normal random sample containing 1,218, 1,198, and 58 data points. The ranges are calculated using a 50,000 repeats Monte Carlo simulation with the same sample sizes as the empirical data sets (i.e. 1,218, 1,198, and 58 data points for the daily, daily recalculated monthly and the monthly returns respectively). Bold numbers in Panel A indicate deviation from the normal distribution based on the distribution of Panel B assuming 95% significance level.

Studentized range is sensitive to deviation from the normal distribution as extreme rates have a crucial impact on the ratio but they are relatively insensitive to asymmetry causing severe type II error (see e.g., Shapiro et al, 1968). Affleck-Graves and McDonald (1989) show that directional tests based on skewness and kurtosis reject more often normality than studentized range does. SR statistics cannot capture the DRMR time series' significant negative skewness (-0.77 on average in our data).

Normality of return data not necessarily induces normality in the disturbances of the market model (see e.g., Affleck-Graves and McDonald, 1989). We test the normality of the market model residuals as well. Table 7 shows that the residuals of the market model on daily returns are non-normal but DRMRs appear to be normal in five cases in our data based on K^2 test. There are more non-normal monthly residuals than return time series (6 vs. 1) in our data.

Normality tests based on kurtosis and skewness and D'Agostino-Pearson K^2 test of residuals

Table 7

Panel A

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
Z(1)	-0.78	-3.68	-0.62	-3.28	-3.93	-8.29	-8.94	-1.15	-5.82	-3.97	-3.98	-3.46	-1.22
p(Z(1))	0.29	0.00	0.33	0.00	0.00	0.00	0.00	0.21	0.00	0.00	0.00	0.00	0.19
Z(2)	5.17	8.65	8.42	8.14	5.08	12.59	11.69	5.36	8.56	6.55	6.64	3.94	4.61
p(Z(2))	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K^2	27.32	88.35	71.21	76.98	41.22	227.25	216.58	30.07	107.18	58.59	59.91	27.48	22.71
$p(K^2)$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
Z(1)	-0.42	0.04	-0.59	4.92	-1.70	-0.80	-6.10	-10.35	-0.37	0.69	0.96	-3.01	-2.28
p(Z(1))	0.36	0.40	0.34	0.00	0.09	0.29	0.00	0.00	0.37	0.31	0.25	0.00	0.03
Z(2)	9.46	7.80	7.11	10.96	10.18	7.57	10.21	12.55	4.77	4.94	7.57	7.21	7.19
p(Z(2))	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K^2	89.68	60.85	50.85	144.32	106.44	57.93	141.53	264.71	22.89	24.92	58.22	60.98	56.88
$p(K^2)$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel B

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
Z(1)	0.90	0.65	1.86	1.40	-0.87	-2.19	-0.97	2.03	-3.42	-0.95	-0.93	-0.87	-0.23
p(Z(1))	0.27	0.32	0.07	0.15	0.27	0.04	0.25	0.05	0.00	0.25	0.26	0.27	0.39
Z(2)	-0.64	-0.27	-0.10	0.36	-0.37	1.04	-0.58	1.38	2.80	0.45	0.43	1.30	-0.74
p(Z(2))	0.33	0.38	0.40	0.37	0.37	0.23	0.34	0.15	0.01	0.36	0.36	0.17	0.30
K^2	1.22	0.50	3.46	2.08	0.89	5.89	1.27	6.04	19.49	1.10	1.04	2.43	0.60
$p(K^2)$	0.54	0.78	0.18	0.35	0.64	0.05	0.53	0.05	0.00	0.58	0.59	0.30	0.74
	14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
Z(1)	-1.92	1.76	0.26	1.84	0.31	2.38	0.26	-3.07	0.59	0.48	-0.03	0.93	2.09
p(Z(1))	0.06	0.08	0.39	0.07	0.38	0.02	0.39	0.00	0.34	0.36	0.40	0.26	0.04
Z(2)	2.24	0.52	-0.50	0.42	0.97	2.46	0.80	2.26	1.71	1.54	0.47	0.64	1.59
p(Z(2))	0.03	0.35	0.35	0.36	0.25	0.02	0.29	0.03	0.09	0.12	0.36	0.33	0.11
K^2	8.71	3.37	0.32	3.57	1.04	11.70	0.71	14.58	3.26	2.60	0.22	1.28	6.91
$p(K^2)$	0.01	0.19	0.85	0.17	0.59	0.00	0.70	0.00	0.20	0.27	0.90	0.53	0.03

Panel C

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
Z(1)	3.52	4.97	5.59	2.45	-0.91	-9.57	-4.94	1.78	-10.83	-6.42	-6.41	2.94	1.78
p(Z(1))	0.00	0.00	0.00	0.02	0.26	0.00	0.00	0.08	0.00	0.00	0.00	0.01	0.08
Z(2)	-1.13	3.68	4.59	3.72	0.68	3.66	0.11	1.08	7.87	5.50	5.46	2.27	-4.49
p(Z(2))	0.21	0.00	0.00	0.00	0.32	0.00	0.40	0.22	0.00	0.00	0.00	0.03	0.00
K^2	13.69	38.28	52.35	19.85	1.30	105.05	24.44	4.35	179.22	71.51	70.82	13.80	23.30
$p(K^2)$	0.00	0.00	0.00	0.00	0.52	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.00
	14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
Z(1)	7.35	2.56	-2.75	5.38	-1.10	0.35	-0.73	-11.96	-0.95	-0.37	-2.85	-4.97	1.43
p(Z(1))	0.00	0.02	0.01	0.00	0.22	0.38	0.31	0.00	0.25	0.37	0.01	0.00	0.14
Z(2)	9.58	2.55	1.19	5.07	0.71	1.32	3.23	7.97	2.78	1.78	-0.02	4.74	3.71
p(Z(2))	0.00	0.02	0.20	0.00	0.31	0.17	0.00	0.00	0.01	0.08	0.40	0.00	0.00
K^2	145.82	13.05	8.95	54.64	1.71	1.87	10.94	206.47	8.65	3.31	8.11	47.14	15.80
$p(K^2)$	0.00	0.00	0.01	0.00	0.43	0.39	0.00	0.00	0.01	0.19	0.02	0.00	0.00

This table shows the test statistics of normality tests and their p-values of market model's residuals using daily, monthly and daily recalculated monthly returns in panel A, B, C respectively. Z(1) tests normality with an alternative hypothesis of non-normal distribution due to skewness. Z(2) tests normality with an alternative hypothesis of non-normal distribution due to kurtosis. D'Agostino-Pearson K2 test statistic is the sum of squares of Z(1) and Z(2) statistics, which tests normality against non-normal distribution due to skewness and kurtosis.

We can look at monthly returns as a sample of DRMRs. Monthly returns compare the last closing price of a given month with the last closing price of the previous month. Assuming 21 trading days a month (which is the case on average) all the monthly returns would be included in the DRMRs population so monthly returns would be a sample of DRMRs. This feature also sheds some light on non-normality of monthly returns. Monthly returns are results of systematic sampling; although, they are not fully random samples since the choice of the sampling starting point is not random.

There are numerous tests for equality of distributions, parametric and non-parametric as well. The time series in our data are not normal and we do not know the exact distribution of the returns, so only non-parametric approaches are applicable.

We use Wilcoxon rank-sum test (Wilcoxon, 1945) which is also known as Mann-Whitney U test or Mann-Whitney-Wilcoxon (henceforth *MWW*) test first introduced by Wilcoxon for equal sample sizes and then extended to arbitrary sample sizes by Mann and Whitney (1947). Hodges and Lehman (1956) argue on theoretical grounds in favor of the rank-sum test, noting that it is never much less efficient than Student's t (where less efficient means more samples are required to get the same performance) but can be infinitely more efficient. Potvin and Roff (1993) conclude that the MWW test is more powerful than the t-test for skewed distributions. MWW test is the best in our case since our data are highly skewed.

MWW test assumes that the two underlying distributions have approximately the same shape (variance) and the only difference between them is a shift in location. 58 observations of DRMRs are chosen five times randomly to obtain five independent random samples. The variances of monthly returns and of random samples from DRMRs are very close to each other indicating a powerful test. We compare the five random samples with the 58 monthly returns using MWW test which can be seen in Table 8 Panel A.

Mann-Whitney-Wilcoxon (MWW) and Kruskall-Wallis (KW) tests

Table 8

Panel A

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
1	1.85	1.50	1.73	1.48	2.00	1.03	0.95	0.12	0.51	1.25	1.25	0.96	0.69
	(0.06)	(0.13)	(0.08)	(0.13)	(0.04)	(0.30)	(0.34)	(0.90)	(0.60)	(0.21)	(0.21)	(0.33)	(0.48)
2	1.40	0.34	0.53	0.58	1.49	0.36	0.27	0.47	0.20	0.74	0.76	1.41	0.42
	(0.16)	(0.73)	(0.59)	(0.56)	(0.13)	(0.71)	(0.78)	(0.63)	(0.84)	(0.46)	(0.44)	(0.15)	(0.67)
3	0.19	0.46	0.09	0.06	0.21	0.46	0.42	0.67	0.18	0.15	0.14	0.79	0.15
	(0.84)	(0.64)	(0.92)	(0.95)	(0.83)	(0.64)	(0.67)	(0.50)	(0.85)	(0.87)	(0.89)	(0.42)	(0.87)
4	0.65	0.09	0.26	0.08	1.13	0.32	0.57	0.35	0.70	0.00	0.01	0.16	0.12
	(0.51)	(0.93)	(0.79)	(0.93)	(0.25)	(0.74)	(0.57)	(0.72)	(0.48)	(0.99)	(0.98)	(0.87)	(0.90)
5	0.91	0.89	0.91	0.85	0.52	1.36	1.50	0.63	1.11	0.74	0.70	0.21	1.00
	(0.36)	(0.37)	(0.36)	(0.39)	(0.60)	(0.17)	(0.13)	(0.52)	(0.26)	(0.46)	(0.48)	(0.83)	(0.31)
	14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
1	0.76	0.80	0.88	0.10	1.14	1.13	0.52	0.64	0.88	0.76	1.02	1.45	0.91
	(0.44)	(0.42)	(0.37)	(0.91)	(0.25)	(0.25)	(0.60)	(0.52)	(0.37)	(0.44)	(0.30)	(0.14)	(0.36)
2	0.90	0.43	0.57	1.02	1.05	0.57	0.00	0.16	0.70	0.82	0.32	0.77	0.70
	(0.36)	(0.66)	(0.57)	(0.30)	(0.29)	(0.56)	(0.99)	(0.87)	(0.48)	(0.41)	(0.74)	(0.44)	(0.48)
3	0.11	0.02	0.43	0.05	0.40	0.01	0.20	0.31	0.05	0.14	0.41	0.68	0.20
	(0.90)	(0.98)	(0.66)	(0.95)	(0.68)	(0.98)	(0.84)	(0.75)	(0.96)	(0.88)	(0.68)	(0.49)	(0.84)
4	0.47	0.55	0.06	0.27	0.63	0.41	0.95	0.12	0.17	0.16	0.47	0.03	0.18
	(0.64)	(0.58)	(0.94)	(0.78)	(0.52)	(0.68)	(0.34)	(0.90)	(0.86)	(0.87)	(0.64)	(0.97)	(0.85)
5	0.78	0.67	1.06	0.22	1.05	0.95	0.82	0.81	0.26	0.23	0.57	0.66	0.64
	(0.43)	(0.50)	(0.29)	(0.82)	(0.29)	(0.34)	(0.41)	(0.41)	(0.79)	(0.81)	(0.56)	(0.50)	(0.52)

Panel B

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
KW	6.76	6.00	4.90	4.28	6.04	4.39	4.89	2.61	1.50	2.88	2.91	3.16	1.43
	(0.24)	(0.31)	(0.43)	(0.51)	(0.3)	(0.49)	(0.43)	(0.76)	(0.91)	(0.72)	(0.71)	(0.68)	(0.92)
	14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
KW	2.13	1.21	2.82	1.29	2.28	2.05	2.24	1.04	2.03	2.01	2.53	4.98	1.68
	(0.83)	(0.94)	(0.73)	(0.94)	(0.81)	(0.84)	(0.81)	(0.96)	(0.84)	(0.85)	(0.77)	(0.42)	(0.89)

This table shows the results of MWW tests comparing the monthly returns with five random samples from daily recalculated monthly returns in Panel A and the results of KW test comparing the monthly returns with the five random samples simultaneously. P-values can be seen in parenthesis.

There is only one case (fund number 5) when the null hypothesis of equality in distributions can be rejected at 95% significance level.

The Kruskal-Wallis (*KW*) test (1952), the generalization of MWW test for more than two samples is also performed (Table 8 Panel B). The KW test confirms the results of MWW tests, the null hypothesis of homogenity of monthly returns and the five random samples cannot be rejected at all usual significance levels. We can conclude that monthly returns and daily recalculated monthly returns have the same distribution, some possible minor differences are only due to random sampling.

We also study the other two essential assumptions in standard asset pricing tests, homoskedasticity and serial independence. Applying the White (1980) test, the market model residuals show significant heteroskedasticity. There are only three mutual funds with heteroskedastic residuals running CAPM regressions on monthly returns. Using DRMRs, we can only accept the homoskedastic null hypothesis in five cases.

Table 9 shows the results of Breusch-Godfrey (1978) autocorrelation tests. Using daily returns, there is significant autocorrelation in the case of 23 mutual funds, while only two mutual funds show autocorrelation in the case of monthly returns at 95% significance level. The market model residuals are heteroskedastic if we use DRMRs.

Heteroskedasticity and autocorrelation tests of the market model residuals

Table 9

$$R_{i,t} = \alpha_i + \beta_i (R_{m,t}) + \varepsilon_{i,t} \quad i = 1,2,...,26; \quad t = 1,2,...,T$$

		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
A	*p(W)*	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	p(BG)	0.04	0.00	0.01	0.03	0.02	0.91	0.04	0.02	0.04	0.00	0.00	0.00	0.00
		14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
	p(W)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	p(BG)	0.00	0.02	0.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.37	0.00
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
B	*p(W)*	0.86	0.79	0.51	0.82	0.91	0.03	0.46	0.52	0.43	0.00	0.00	0.78	0.69
	p(BG)	0.10	0.30	0.29	0.46	0.50	0.35	0.09	0.78	0.97	0.98	0.98	0.18	0.07
		14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
	p(W)	0.84	0.26	0.85	0.20	0.85	0.85	0.31	0.16	0.15	0.30	0.91	0.27	0.57
	p(BG)	0.00	0.35	0.85	0.29	0.93	0.81	0.00	0.19	0.40	0.65	0.14	0.51	0.88
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
C	*p(W)*	0.51	0.03	0.02	0.03	0.32	0.18	0.00	0.00	0.44	0.00	0.00	0.18	0.00
	p(BG)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		14	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*
	p(W)	0.00	0.00	0.04	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00
	p(BG)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

This table shows p-values of the White heteroskedasticity (p(W)) and Breusch-Godfrey autocorrelation tests (p(BG)) on the market model's residuals using daily, monthly, and daily recalculated monthly returns in panel A, B, C respectively. We use the Morgan Stanley Capital International (MSCI) All Countries World Index as the market proxy.

Heteroskedasticity and autocorrelation make the standard errors of *OLS* (Ordinary Least Squares) parameter estimations inconsistent and biased so the usual *t-* and *F-tests* cannot be applied. If we use *OLS*, there are methods generating consistent covariance matrix even in the case of heteroskedasticity, such as Heteroskedasticity Consistent Covariance Matrix (*HCCM*) estimation (White, 1980). If the residuals are heteroskedastic and autocorrelated at the same time, the Heteroskedasticity and Autocorrelation Consistent (*HAC*) covariance matrix estimation (Newey and West, 1987, 1994; Andrews, 1991; Andrews and

Monahan, 1992) combined with OLS parameter estimations are appropriate to make the usual F- and t-tests applicable. Therefore we calculate t-statistics with HAC covariance matrix (Andrews, 1991).

If continuously compounded daily returns are independent and identically distributed (*i.i.d.*), then on the basis of central limit theorem by increasing the number of periods to be summed, the distribution of cumulated returns is approaching the normal distribution. If the central limit theorem is valid, the distribution of monthly returns should be more normal than the daily ones. However, independence does not exist in the case of daily returns (Table 9). Furthermore, Mandelbrot (1963) proves that the central limit theorem is a special case and it is valid only if variance of the random variable is finite. Generally the distribution converges to a stable but not necessarily the normal distribution. Not normal stable symmetrical distributions have a leptokurtic feature which can be observed in daily returns (kurtosis is 5,47 on average in our data).

As a result of stable distribution, if daily returns are not normally distributed with infinite variance, then distribution of weekly and monthly returns will not be normal, irrespectively of the length of summation.

Fama (1976) shows empirically that the distribution of monthly returns is less leptokurtic and the frequency of extreme returns is smaller than in the case of daily returns. Unlike the theorem of stable distribution, the distribution of monthly returns approach the Gaussian distribution rather than the distribution of daily returns, but relative frequency of extreme returns and returns close to the average turn out to be a bit higher than that of normal distribution. Although our tests show that monthly returns only seem to be normal because of the too short time series and severe type II error.

Blume (1968) and Officer (1971) thoroughly study returns of single securities and portfolios and find that distribution of portfolios is of the same type as that of single securities, so it does not matter if we examine single securities or as, at present, mutual funds (portfolios).

4. Visualizing non-normality

Distribution of monthly returns and DRMRs of two mutual funds can be seen in Figures 4-5. In the case of DRMRs the fat tail phenomenon can be observed on the left of the distribution, i.e. big losses are more frequent than it would be drawn from normal distribution. Small positive returns are more likely than they would be in the case of Gaussian distribution, which is well presented by the negative skewness of the distributions as the -0.87% and -0.43% in the case of mutual fund number 25 and 12 respectively. Return distribution of fund number 25 shows leptokurtosis contrary to the normal distribution (kurtosis 4.5), distribution of fund number 12 does not significantly differ from Gaussian distribution from the kurtosis point of view (3.18).

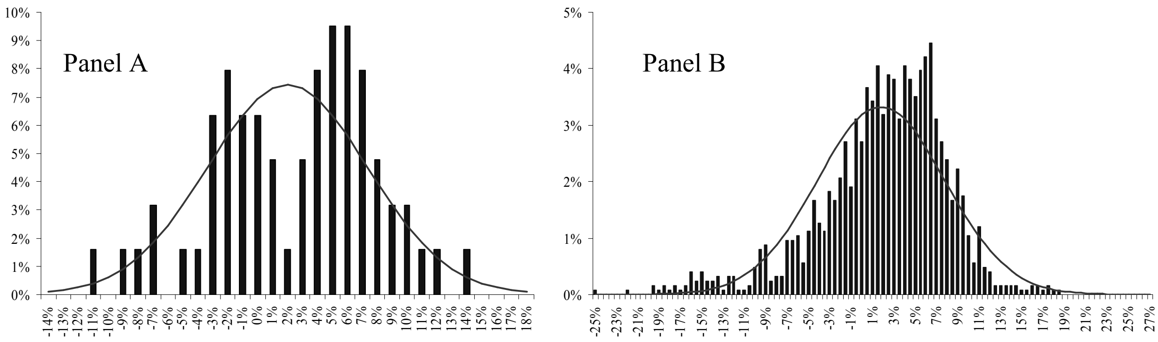

Figure 4 Histograms of monthly returns and daily recalculated monthly returns (*DRMR*s) of mutual fund number 25.

Notes: Histograms of monthly returns and *DRMR*s can be seen in Panel A and Panel B respectively. Fund number 25 is chosen for illustration because its $\boldsymbol{R^2}$ is equal to the average of the 26 market model regressions included in our paper using when *DRMR*s. Red curves show the normal probability density functions.

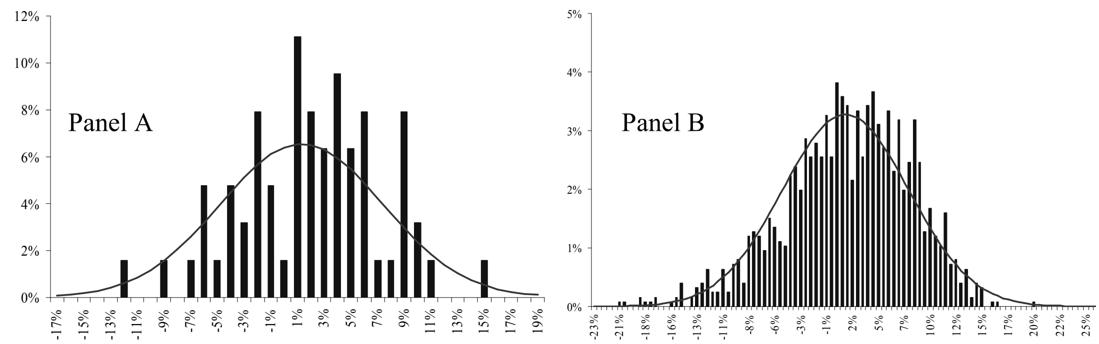

Figure 5 Histograms of monthly returns and daily recalculated monthly returns (*DRMR*s) of mutual fund number 12.

Notes: Histograms of monthly returns and *DRMR*s can be seen in Panel A and Panel B respectively. Fund number 12 is chosen for illustration because its $\boldsymbol{R^2}$ is equal to the average of the 26 market model regressions included in our paper when using daily returns. Red curves show the normal probability density functions.

Fat tail can also be observed in the case of simple monthly returns, i.e. on the sides of empirical distribution events are more frequent than it would be drawn from normal distribution.

Looking at Figure 4-5, monthly returns are not normally distributed confirming our results in the previous section. The only reason why they seem to be normal when they are tested by D'Agostino-Pearson or other statistics is the small sample size.

If we intend to study the deviation of the empirical distribution from the normal, comparing the logarithm of the density functions can be an illustrative solution. This procedure highlights anomalies of the empirical distribution like fat tails. The logarithm of a normal probability density function is a parabolic curve. The use of this method is practical because the logarithm enlarges the difference between the empirical and the normal distribution most significantly in the tails.

According to Figures 6-7 daily recalculated returns fit the normal distribution better than monthly returns in the centre of the distribution. DRMRs of fund number 25 exhibit significant departures from Gauss-distribution at the tails, while in the case of fund number 12 there is no significant deviation either in the centre or at the tails.

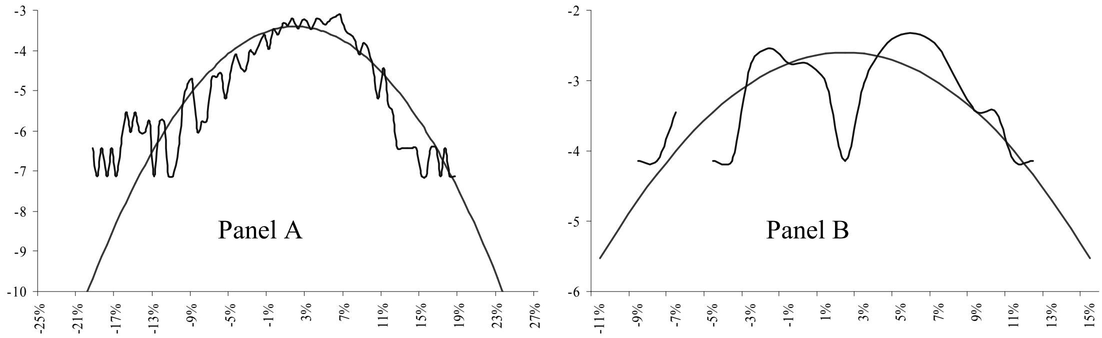

Figure 6 Logarithm of density function of daily recalculated monthly returns (*DRMR*s) and monthly returns of mutual fund number 25

Notes: Logarithm of density function of *DRMR*s and monthly returns can be seen in Panel A and Panel B respectively. Fund number 25 is chosen for illustration because its $\boldsymbol{R^2}$ is equal to the average of the 26 market model regressions included in our paper when using *DRMR*s. Red curves show the logarithm of normal probability density functions.

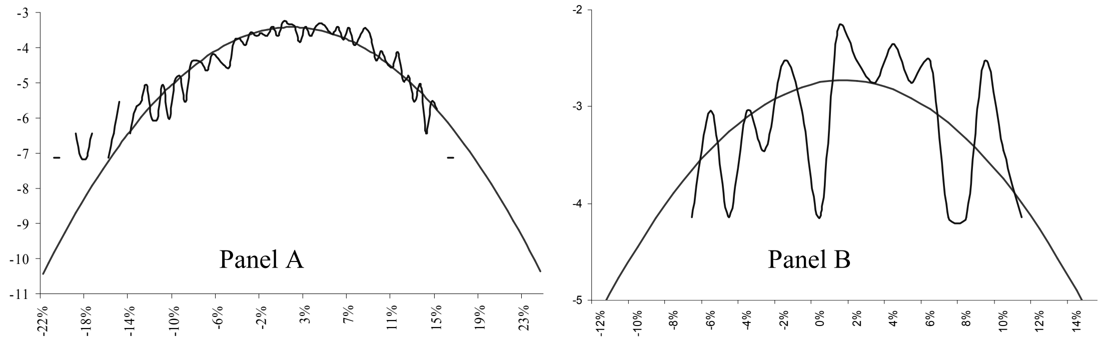

Figure 7 Logarithm of density function of *DRMR*s and monthly returns of mutual fund number 12

Notes: Logarithm of density function of *DRMR*s and monthly returns can be seen in Panel A and Panel B respectively. Fund number 12 is chosen for illustration because its $\boldsymbol{R^2}$ is equal to the average of the 26 market model regressions included in our paper when using *DRMR*s. Red curves show the logarithm of normal probability density functions.

Based on the logarithms of empirical density functions we cannot say that monthly returns are more normal than DRMRs.

5. Concluding remarks

When mutual funds performance is to be evaluated, or the efficiency of capital market is to be tested, the key question is whether there is any significant abnormal return. The method used most frequently to settle the question is the calculation of Jensen alphas and their univariate or multivariate (GRS test) econometric study. From the returns calculation point of view, we face two issues. Firstly, adequate sample size is required for reliable parameter estimations. Secondly, the fewer assumptions are needed in connection with the distribution, the larger sample size is required. With large sample size, asymptotic features of the estimations can be derived, i.e. consistent standard errors of parameters which are essential for performance measurement.

In the literature most frequently daily or monthly returns are used and there is no special recipe for measuring debuting assets' performance or testing market efficiency of newly traded markets. In this paper we examine an alternative return calculation, the daily recalculated monthly returns (DRMR). Although, DRMRs are not normally distributed, serially dependent, and heteroskedastic, they can be used when no long enough time series is available. If we work on short time series, monthly returns are not applicable because the parameter estimations are not reliable and we cannot derive asymptotic asset pricing tests when, for example, normality is not valid. We show that monthly returns are not normally distributed; although, they seem to be so in small samples, but they are a systematic sample of the DRMR population.

Another possibility to solve our problem is the daily return calculation. Daily returns have all the bad properties that DRMRs do and they are the most sensitive to measurement errors in the market proxy resulting in weak fitting.

We show that DRMRs are the best working approximation when we have very limited data available. This method gives at least as good fitting of the market model as the monthly returns or even better (relatively high R^2). At the same time it does not imply dramatic decrease of data points suitable for asymptotic tests even for relatively short periods.

6. References

1. Affleck-Graves, J., and B. McDonald. "Nonnormalities and Tests of Asset Pricing Theories." *The Journal of Finance,* 44 (1989), 889-908.
2. Andrews, D.W.K. Heteroskedasticity and "Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica,* 59 (1991), 817-858.
3. Andrews, D.W.K., and J.C. Monahan, "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator." *Econometrica,* 60 (1992), 953-966.
4. Black, F.; M.C. Jensen; and M. Scholes. "The Capital Asset Pricing Model: Some Empirical Tests." In: Jensen M. C. (Eds.), Studies in the Theory of Capital Markets, Praeger, New York, (1972). 79-121.
5. Blume, M. E. "The Assessment of Portfolio Performance." Unpublished Ph.D. dissertation. University of Chicago (1968).
6. Breusch, T. "Testing for Autocorrelation in Dynamic Linear Models." *Australian Economic Papers,* 17 (1978), 334-55.
7. Brown, S. J., and J. B. Warner. "Using Daily Stock Returns: The Case of Event Studies." *Journal of Financial Economics,* 14 (1985), 3-31.
8. Chen, H., and V. Singal. "Role of Speculative Short Sales in Price Formation: The Case of the Weekend Effect". *The Journal of Finance*, 58 (2003), 685-705.
9. Cochrane, J. H. "Asset Pricing". Princeton University Press, Princeton (2001).
10. D'Agostino, R.B; A. Belanger; and R.B. D'Agostino Jr. "A Suggestion for Using Powerful and Informative Tests of Normality". *The American Statistician,* 44 (1990), 316-321.
11. De Santis, G., and B. Gerard. "International Asset Pricing and Portfolio Diversification with Time-varying Risk." *The Journal of Finance*, 52 (1997), 1881-1912.
12. Dickey, D.A., and W.A. Fuller. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root." *Econometrica,* 49 (1981), 1057-1072.

13. Elton, E.J.; M.J. Gruber; and C.R. Blake. "Fundamental Economic Variables, Expected Returns, and Bond Fund Performance." *The Journal of Finance*, 50 (1995), 1229-1256.
14. Elton, E.J.; M.J. Gruber; and C.R. Blake. "Survivorship Bias and Mutual Fund Performance." *The Review of Financial Studies,* 9 (1996), 1097-1120.
15. Fama, E.F., and J. D. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *The Journal of Political Economy*, 81 (1973), 607-636.
16. Fama, E.F. Foundations of Finance, Basic Books, New York (1976).
17. Fama, E.F., and K.R. French. "Size and Book-to-market Factors in Earnings and Returns." *The Journal of Finance*, 50 (1995), 131-155.
18. Fama, E.F., and K.R. French. "Multifactor Explanations of Asset Pricing Anomalies." *The Journal of Finance,* 51 (1996a), 55-84.
19. Fama, E.F., and K.R. French. "The CAPM is Wanted, Dead or Alive." *The Journal of Finance*, 51 (1996b), 1947-1958.
20. Fama, E.F., and K.R. French. "Value Versus Growth: The International Evidence." *The Journal of Finance*, 53 (1998), 1975-1999.
21. Fama, E.F., and K.R. French. "The Value Premium and the CAPM." *The Journal of Finance*, 61 (2006), 2163-2185.
22. Gibbons, M.R.; S.A. Ross; and J. Shanken. "A Test of the Efficiency of a Given Portfolio." *Econometrica*, 57 (1989), 1121-1152.
23. Godfrey, L. G. "Testing for Higher Order Serial Correlation in Regression Equations when the Regressors Include Lagged Dependent Variables." *Econometrica*, 46 (1978), 1303-1310.
24. Jarque, C.M., and A.K. Bera. "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals." *Economics Letters,* 6 (1980), 255–259.
25. Jensen, M.C. "The Performance of Mutual Funds in the Period 1945-1964." *The Journal of Finance*, 23 (1968), 389-416.
26. Hansen, L.P. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica*, 50 (1982), 1029-1054.
27. Hodges, J.L. Jr, and E. L. Lehmann, "The Efficiency of Some Nonparametric Competitors of the "t"-test." *The Annals of Mathematical Statistics,* 27 (1956), 324-335.
28. Kothari, S.P.; J.Shanken; and R.G. Sloan. "Another Look at the Cross-section of Expected Stock Returns." *The Journal of Finance,* 50 (1995), 185-224.
29. Kruskal, W.H., and W.A. Wallis. "Use of Ranks in One-criterion Variance Analysis." *Journal of the American Statistical Association*, 47 (1952), 583-621.
30. Mandelbrot, B. "The Variation of Certain Speculative Prices." *The Journal of Business,* 36 (1963), 94-419.
31. Mann, H.B., and D.R. Whitney. "On a Test of whether One of Two Random Variables is Stochastically Larger than the Other." *The Annals of Mathematical Statistics*, 18 (1947), 50-60.
32. Morse, D. "An Econometric Analysis of the Choice of Daily Versus Monthly Returns in Tests of Information Content." *Journal of Accounting Research*, 22 (1984), 605-623.
33. Newey, W.K., and K.D. West. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica,* 55 (1987), 703-708.
34. Newey, W.K., and K.D. West. "Automatic Lag Selection in Covariance Matrix Estimation." *The Review of Economic Studies,* 61 (1994), 631-653.

35. Officer, R. "A Time Series Examination of the Market Factor of the New York Stock Exchange." Ph.D. dissertation, University of Chicago (1971).
36. Potvin, C., and D.A. Roff. "Distribution-free and Robust Statistical Methods: Viable Alternatives to Parametric Statistics." *Ecology*, 74 (1993), 1617-1628.
37. Roll, R. "A Critique of the Asset Pricing Theory's Tests' Part I: On Past and Potential Testability of the Theory." *Journal of Financial Economics*, 4 (1977), 129-176.
38. Shapiro, S.S.; M.B. Wilk; and H.J. Chen. "A Comparative Study of Various Tests for Normality." *Journal of the American Statistical Association*, 63 (1968), 1343-1372.
39. Sharpe, W.F. "Mutual Fund Performance." *The Journal of Business,* 39 (1966), 119-138.
40. Solnik, B. "The International Pricing of Risk: An Empirical Investigation of the World Capital Market Structure." *The Journal of Finance*, 29 (1974), 365-378.
41. White, H. "A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica*, 48 (1980), 817-838.
42. Wilcoxon, F. "Individual Comparisons by Ranking Methods." *Biometrics Bulletin*, 1 (1945), 80-83.